Endopeptidase Based Metabolic Pathway Engineering In *Cryptococcus neoformans*

Pradeep Kumar Mohapatra, Bibhuti Prasad Barik

Department of Zoology, Khallikote Unitary University, Brahmapur, Odisha

Abstract

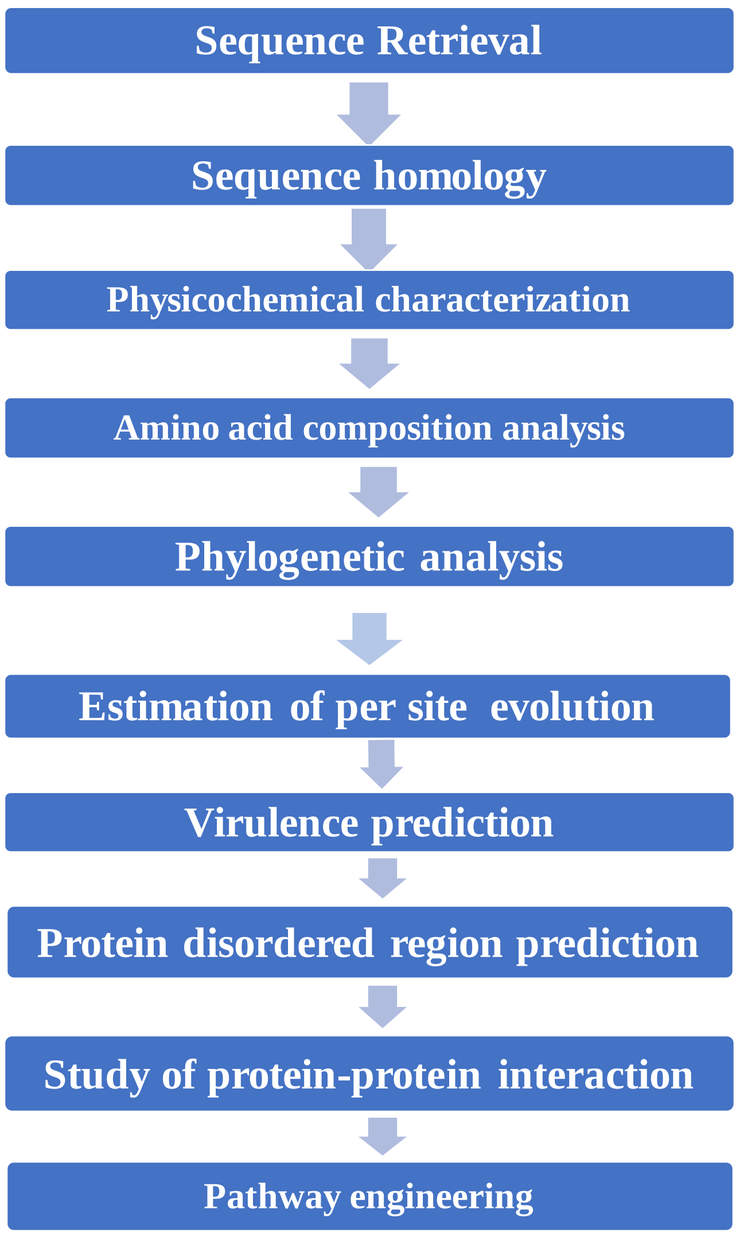
Cryptococcus neoformans is a fungal pathogen causing cryptococcosis. It primarily infects lungs & brain tissues with symptoms like cough & chest pain leading to death in immunosuppressed individuals. Aspartic type endopeptidase (AEP) found in the fungal capsule plays a crucial role in its virulence. This study characterizes the AEP protein in silico & proposes a putative pathway to understand its mechanism of action. Physicochemical characterization, virulence prediction & protein-protein interaction studies were carried out indicating similarities with other homologous species. The physicochemical characterization of AEP in different fungal species suggested the optimum pH for the activity of this enzyme as well as in measuring the thermostability of the proteins & calculated hydrophobicity of the protein highlighted the functionality of the protein. Disordered region prediction explains their catalytic & infectious role. A putative pathway was designed for the mode of action of this protein by analysing protein-protein interaction data & pathway data from different databases which predicts the most effective pathway for the protein. The proposed pathway may highlight on virulence properties of AEP & potential drug development against *C. neoformans* infection.

Objectives:

This study has following objectives through which we want to decipher the fungal protein & to generate a comprehensive idea about the nature, structure, function, interaction & evolutionary significance of the protein.

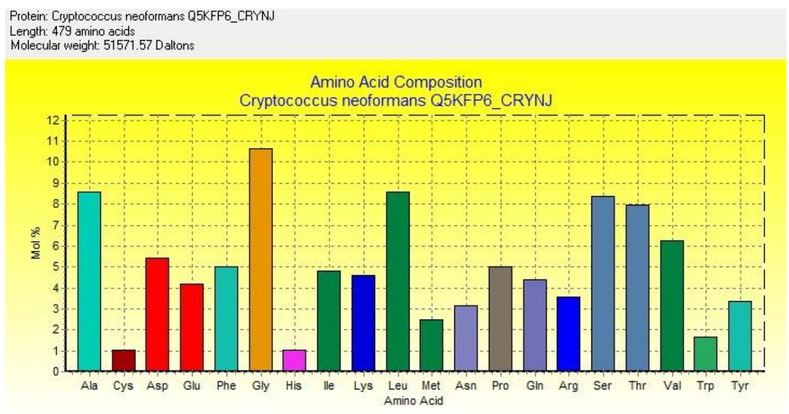
- *In silico* analysis & prediction of virulence of the proteins
- Prediction of disordered region of these proteins for studying protein - protein interactions, protein function, & potential drug designing against the proteins
- Prediction of protein – protein interactions
- Establishing phylogenetic relationships between the homologous proteins
- Putative pathway prediction for the target protein

Materials & Methods:



Result & Discussion:

The physicochemical characterization of the protein revealed that the all the twenty protein considered in this study have similar molecular weight suggesting their similar catalytic efficiency (Wilk et al., 1983). The TpI value have ranged from 4.37 to 6.65 showing that the enzymes are active in different physiological pH (Gasteiger et al., 2005). The I.I of most of the proteins are below 40 signifying that they are having a structural stability (Jin et al.,2016). The A.I value suggested that all the 20 proteins are thermostable as they have higher A.I value. GRAVY value highlights that out of all the considered proteins 7 are hydrophobic as they have positive GRAVY value.


Protein: Cryptococcus neoformans Q5KFP6_CRYNJ
Length: 479 amino acids
Molecular weight: 51571.57 Daltons
Amino Acid Composition
Cryptococcus neoformans Q5KFP6_CRYNJ
Mol %
Ala Cys Asp Glu Phe Gly His Ile Lys Leu Met Asn Pro Gln Arg Ser Thr Val Trp Tyr
Amino Acid

The target protein have highest content of Glycine & alanine which accompany the protein to interact with a wide range of substrate (Tsumoto et al., 2004). The phylogenetic analysis of the target & its homolog resulted into a tree which shows that species belonging to *Cryptococcus* genus are evolutionary related to each other & diverged from each other in the recent past. *Kwoniella* species also shows stronger evolutionary relationship as they resides in a group in the phylogenetic tree. The tree also resulted into a separate lineage for *Trichosporona* & *Tulosesus* genus showing they have diverged from the ancestral lineage much earlier than *Cryptococcus* & *Kwoniella* genus.

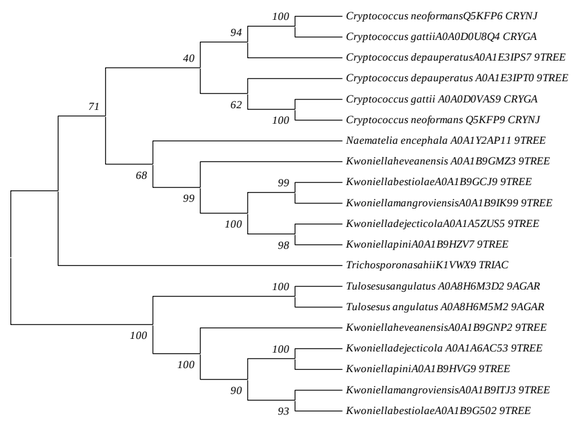
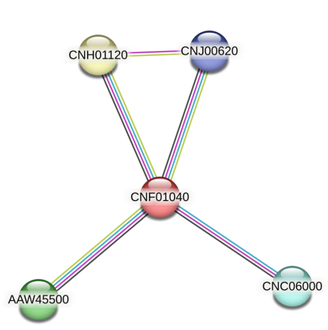


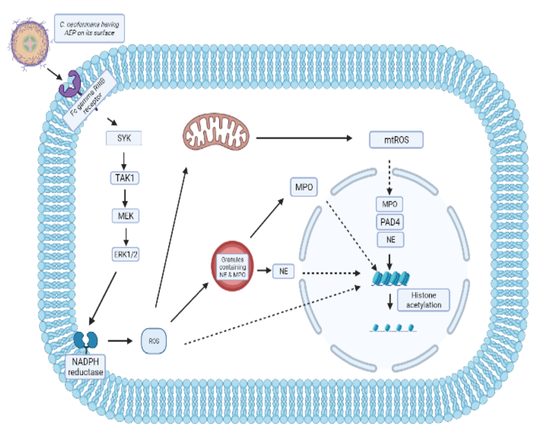
Fig: Phylogenetic tree obtained using maximum likelihood method

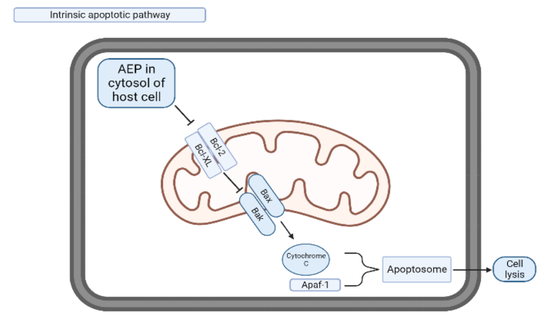
The per site evolutionary rate prediction shows that the protein have most conserved sequence between 71-80, 88-106, 141-151, 295-310 & 391-415 amino acid sequences hence these regions of protein may be taking part in catalytic activity of the protein. Similarly, virulence prediction showed that out of 29 considered protein 19 are having virulence potential hence endopeptidase in different class of fungi have pathogenic potential (Garg et al., 2008). The protein disordered region prediction produced results which shows all the proteins have disordered region in them which may be helping the protein to interact with its host cell & substrates (Uversky et al., 2015).

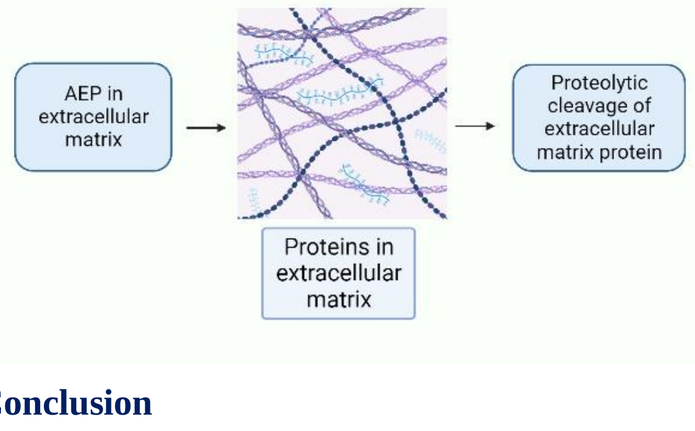


The conservation scale:
1 2 3 4 5 6 7 8 9
Variable
Average
Conserved
e - An exposed residue according to the NACSES algorithm.
b - A buried residue according to the NACSES algorithm.
f - A predicted functional residue (highly conserved and exposed).

The protein-protein interaction data from STRING shows that the target protein mainly interact with three other proteins namely Kex protein, Actin cross linking protein & serine type endopeptidase showing that the protein plays different role like catalytic activator, actin depolymerizer by interacting with these proteins (Tseng et al., 2005). The pathway analysis led to proposal of mainly three pathways for mechanism of action of the protein in different environmental conditions which are as follows:






AEP in extracellular matrix
Proteolytic cleavage of extracellular matrix protein
Proteins in extracellular matrix

Conclusion

This study may lead to predict new function, physicochemical properties, active sites for drug designing as well as to study evolutionary relationship between homologous proteins hence providing a knowledge about the evolutionary history of the species & how it has evolved from the ancestral lineage. The predicted disordered region of the protein will provide a great source for drug targeting. The per site evolutionary rate may be useful in determining the highly conserved regions of this protein thus providing the knowledge about the active sites of the protein. The proposed pathway can lead to designing of potential drugs for effective control of the pathogenicity of the protein hence the organism.